

Optimized File Cache Organization in a Network Server

ABSTRACT

5 A data processing server and method in which the server device stores a first fragment of
a requested file in a first tier of storage while retaining subsequent fragments of the file in a
lower tier of storage. The first tier is typically the server's volatile system memory while the
second tier may represent a local disk, a networked storage device, or a remote system memory.
When the server receives a client request for a file, the server transmits a first fragment of the file
10 stored in the file cache to the client. Simultaneously, the server retrieves a subsequent fragment
of the file from a lower tier of storage. By the time the first fragment is transmitted and
acknowledged, the subsequent fragment is ready for transmission. In this manner, the server is
able to maintain responsiveness while minimizing the amount of data cached in valuable system
memory.

T02260 60059660